



MM-20108 (135224-1)

PATENT

AF
IHW

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES**

Applicant(s):	Hoashi et al.)	Group Art Unit: 2142
)	
Serial No.:	09/503,757)	Examiner: Hai V. Nguyen
)	
Filed:	February 14, 2000)	
)	
For:	METHOD AND APPARATUS FOR AUTOMATIC)	
	INFORMATION FILTERING USING URL)	
	HIERARCHICAL STRUCTURE AND)	
	AUTOMATIC WORD WEIGHT LEARNING)	

LETTER

Mail Stop APPEAL BRIEF-PATENTS
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

In response to the Office Action mailed November 4, 2004, applicants enclose three (3) copies of a revised Brief on Appeal. The revised Brief on Appeal includes on pages 2-5 a concise explanation of the claimed invention that refers to the specification by page and line number and to the drawing by reference characters and is believed to comply fully with 37 CFR 1.192(c). The subject matter on pages 6-24 of the revised Brief on Appeal is unchanged from the previous submission.

Respectfully submitted,

Date: November 24, 2004

By: _____

Seymour Rothstein (Reg. No. 19,369)

OLSON & HIERL, LTD.
20 North Wacker Drive - Suite 3600
Chicago, IL 60606
(312) 580-1180

CERTIFICATE OF MAILING

I hereby certify that this paper and its attachments are being deposited with the United States Postal Service on the date shown below with sufficient postage as First Class Mail in an envelope addressed to: Mail Stop APPEAL BRIEF-PATENTS, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450 on November 24, 2004.

Seymour Rothstein



135224-1

PATENT

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES**

Applicant(s):	Hoashi et al.)	Group Art Unit: 2142
)	
Serial No.:	09/503,757)	Examiner: Hai V. Nguyen
)	
Filed:	February 14, 2000)	
)	
For:	METHOD AND APPARATUS FOR)	
	AUTOMATIC INFORMATION)	
	FILTERING USING URL)	
	HIERARCHICAL STRUCTURE AND)	
	AUTOMATIC WORD WEIGHT)	
	LEARNING)	

BRIEF ON APPEAL

Mail Stop APPEAL BRIEF-PATENTS
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

This is an appeal from the final rejection of claims 1-15 in the above-identified application. The original brief was submitted in triplicate and was accompanied by Check No.026322 in the amount of \$330 to cover the fee required under 37 CFR Sec. 1.17(c). Three copies of this revised Brief on Appeal are enclosed.

1. Real Party in Interest

This application is assigned to KDD Corporation

2. Related Appeals and Interferences

There are no related appeals or interferences.

3. Status of All Claims

Claims 1-15, inclusive, are on appeal. Claims 1-15 are original claims and have not been amended during prosecution. The claims on appeal are set forth in Appendix 1 hereto for ready reference.

4. Status of All Amendments Filed After Final Rejection

An Amendment After Final Rejection, filed on or about May 10, 2004, was considered by the Examiner, but he indicated in an Advisory Action dated July 19, 2004 that it did not place the application in condition for allowance.

5. Concise Summary of the Invention

The present invention provides method and apparatus for automatic information filtering which are capable of improving both precision and recall and accurately judging inappropriateness of the content even for a page that contains very few or no text information and only displays images, by utilizing an upper level URL of a URL given in a hierarchical structure.

In one aspect, the invention is a method of automatic information filtering for identifying inappropriate information among various information provided through Internet and blocking presentation of identified inappropriate information. The method comprises the steps of: entering an HTML (HyperText Markup Language) information provided through the Internet (Fig. 6, step S11, and specification page 18, lines 15-30); judging whether a URL (Uniform Resource Locator) of said HTML information entered from the Internet is a top page URL or not, the top page URL being a URL ending with a prescribed character string defining according to a URL hierarchical structure by which each URL is constructed (Fig. 6, step S13, and specification

page 18, line 25 - page 19, line 2); extracting words appearing in information indicated by the top page URL and carrying out an automatic filtering to judge whether said information indicated by the top page URL is inappropriate or not according to the words extracted from said information indicated by the top page URL, when said URL of said HTML information is the top page URL; (Fig. 6, step S15, and specification page 18, line 30 - page 19, line 2); registering an upper level URL derived from the top page URL into an inappropriate upper level URL list and blocking presentation of said information indicated by the top page URL, when said information indicated by the top page URL is judged as inappropriate by the automatic filtering, the upper level URL being derived from the top page URL by keeping a character string constituting the top page URL only up to a rightmost slash (Fig. 6, steps S21 and S31, and specification page 19, lines 3-10); comparing said URL of said HTML information with each URL registered in the inappropriate upper level URL list and judging whether there is any matching URL in the inappropriate upper level URL list when said URL of said HTML information is not the top page URL, and blocking presentation of information indicated by said URL of said HTML information when there is a matching URL in the inappropriate upper level URL list, the matching URL being one upper level URL whose character string is contained in said URL of said HTML information (Fig. 6 and specification page 19, lines 15-25); extracting words appearing in said information indicated by said URL of said HTML information, and carrying out the automatic filtering to judge whether said information indicated by said URL of said HTML information is inappropriate or not according to the words extracted from said information indicated by said URL of said HTML information, when there is no matching URL

in the inappropriate upper level URL list (Fig. 6, steps S25, S27, S29, and specification page 19, lines 26-33); and blocking presentation of said information indicated by said URL of said HTML information when said information indicated by said URL of said HTML information is judged as inappropriate by the automatic filtering (Fig. 6, step S31, and specification page 19, lines 34-page 20 line 2).

In another aspect, the invention pertains, more particularly, to an automatic information filtering apparatus for identifying inappropriate information among various information provided through Internet and blocking presentation of identified inappropriate information. The apparatus comprises an input unit 1 for entering an HTML (HyperText Markup Language) information provided through the Internet (Fig. 5 and specification page 16, lines 13-15). The apparatus includes a top page URL judging unit for judging whether a URL (Uniform Resource Locator) of said HTML information entered from the Internet is a top page URL or not, the top page URL ending with a prescribed character string defined according to a URL hierarchical structure by which each URL is constructed (Fig. 6 and specification page 19, lines 25-30). A first automatic filtering unit 9 is provided for extracting words appearing in information indicated by the top page URL and carrying out an automatic filtering to judge whether said information indicated by the top page URL is inappropriate or not according to the words extracted from said information indicated by the top page URL, when said URL of said HTML information is the top page URL, (Fig. 5 and specification page 16 lines 21-22, and Fig. 6 and specification page 18 line 25 - page 19, line 2). The apparatus includes an inappropriate upper level URL list registration unit for registering an upper level URL derived from the top page URL into an

inappropriate upper level URL list and blocking presentation of said information indicated by the top page URL is judged as inappropriate by the automatic filtering, the upper level URL being derived from the top page URL by keeping a character string constituting the top page URL only up to a rightmost slash (Fig. 5 and specification page 16, lines 21-22, and Fig. 6 and specification page 18, line 25 - page 19, line 21). Further, the apparatus includes an inappropriate URL judging unit for comparing said URL of said HTML information with each URL registered in the inappropriate upper level URL list and judging whether there is any matching URL in the inappropriate upper level URL list when said URL of said HTML information is not the top page URL, and blocking presentation of information indicated by said URL of said HTML information when there is matching URL in the inappropriate upper level URL list, the matching URL being one upper level URL whose character string is contained in said URL of said HTML information (Fig. 6 and specification page 19, lines 15-25). Also, the apparatus includes a second automatic filtering unit 3 for extracting words appearing in said information indicated by said URL of said HTML information, and carrying out the automatic filtering to judge whether said information indicated by said URL of said HTML information is inappropriate or not according to the words extracted from said information indicated by said URL of said HTML information, when there is no matching URL in the inappropriate upper level URL list (Fig. 6 and specification page 19, lines 26-33); and an information presentation blocking unit for blocking presentation of said information indicated by said URL of said HTML information is judged as inappropriate by the automatic filtering (Fig. 6 and specification page 19, lines 34 - page 20, line 6).

6. Concise Statement of All Issues Presented for Review

The following issues are presented for review:

1. Whether claims 1-15, inclusive, are anticipated by Hughes et al. under 35 U.S.C. 102.

7. Grouping of Claims for Each Ground of Rejection

Since all of claims 1-15 were rejected as being anticipated by Hughes et al., all claims stand or fall together.

Claims 1 and 2 pertain to a method of automatic information filtering for identifying inappropriate information among various information provided through Internet and blocking presentation of identified inappropriate information.

Claims 3 and 4 relate to automatic information filtering apparatus for identifying inappropriate information among various information provided through Internet and blocking presentation of identified inappropriate information.

Claims 5 and 6 are directed to a computer usable medium having computer readable program codes embodied therein for causing a computer to function as an automatic information filtering apparatus for identifying inappropriate information among various information provided through Internet and blocking presentation of identified inappropriate information.

Claims 7, 8, and 9 are directed to a method of automatic information filtering and claims 10, 11, and 12 are directed to a parallel automatic information filtering apparatus.

Claims 13-15 pertain to a computer usable medium.

8. Argument**Claims 1-15 are not anticipated by Hughes et al. (U.S. Patent No. 6,065,055)**

While the basis for the final rejection is that claims 1-15 are “anticipated by Hughes et al. (U.S. Patent No. 6,065,055)”, see page 2 of the Final Rejection dated November 14, 2003, it is observed that in the Response to Arguments in the Final Rejection, the Examiner referred to Russell-Falla et al. (U.S. Patent No. 6,266,664) and Humes (U.S. Patent No. 5,996,011) on page 8 of the Final Rejection dated November 14, 2003 to buttress his argument. Applicants submit that these secondary references should be ignored since they are not part of the final rejection. The sole basis of rejecting claims 1-15 in the Final Rejection is that they were anticipated by Hughes et al. The sole basis for the rejection is anticipation under 35 U.S.C. 102(e), and not obviousness under 35 U.S.C. 103.

The Manual of Patent Examining Practice is clear that “for anticipation under 35 U.S.C. 102, the reference must teach every aspect of the claimed invention either explicitly or implicitly. Any feature not directly taught must be inherently present.” MPEP 706.02.

The Examiner bears the burden of demonstrating that each and every claimed element is present in the anticipatory reference. Anticipation is not established if in reading a claim on something disclosed in a reference it is necessary to pick, choose and combine various positions of the disclosure not directly related to each other by the teachings of the reference. Ex. parte Beuther, 71 USPQ 2d 1313 (Bd of Pat. App. & Int. 2003).

The courts take a stringent view of anticipation as well. A patent is invalid when the same device or method, having all of the elements and limitations contained in the claims, is described in a single prior art reference. ATD Corp. v. Lydall, Inc., 48 USPQ 2d 1321, 1328 (Fed. Cir. 1998). Invalidity for anticipation requires that all of the elements and limitations of the claim are found within a single prior art reference. Scripps Clinic v. Genentech Inc., 18 USPQ 2d 1001, 1010 (Fed. Cir. 1991). See also General Electric Co. v. Nintendo Co. Ltd., 50 USPQ 2d 1910 (Fed. Cir. 1999). A prior art reference anticipates a claim only if the reference discloses, either expressly or inherently, every limitation of the claim. Absence from the reference of any claimed element negates anticipation. Rowe v. Dror, 42 USPQ 2d 1550, 1553 (Fed. Cir. 1997). Such test is strictly applied.

Applicants submit that claims 1-15 patentably distinguish over Hughes et al. and are not anticipated by Hughes et al.

Turning first to claim 1, it should be noted that, in the automatic information filtering method of claim 1, whether the URL of the entered HTML information is a top page URL or not is judged first. If it is a top page URL, words appearing in the information indicated by that top page URL (i.e., not that URL itself but the information provided at that URL) are extracted and the automatic filtering to judge whether this information is inappropriate or not is carried out. If it is judged as inappropriate, an upper level URL derived from that top page URL is registered into an inappropriate upper level URL list and the presentation of the information indicated by that top page URL is blocked. On the other hand, if the URL of the entered HTML information is not a top page URL, this URL is compared with each URL registered in the inappropriate

upper level URL list, and if there is a matching URL in the inappropriate upper level URL list, the presentation of information indicated by this URL blocked, whereas if there is no matching URL in the inappropriate upper level URL list, words appearing in the information indicated by that URL (i.e., not that URL itself but the information provided at that URL) are extracted and the automatic filtering to judge whether this information is inappropriate or not is carried out. Then, if it is judged as inappropriate, the presentation of the information indicated by that URL is blocked.

In other words, when the URL of the entered HTML information is either a top page URL or a non-top page URL with no matching URL registered in the inappropriate upper level URL list, the automatic filtering according to words extracted from the information indicated by that URL is carried out to determine whether the presentation of the information indicated by that URL should be blocked or not. On the other hand, when the URL of the entered HTML information is a non-top page URL of the entered HTML information is a non-top page URL with a matching URL registered in the inappropriate upper level URL list, the presentation of the information indicated by the URL is immediately blocked without carrying out the automatic filtering. Here, the upper level URLs registered in the inappropriate upper level URL list are derived from the top page URLs whose information was found inappropriate in the past.

In this way, it becomes possible to judge the inappropriateness of the HTML information accurately even in the case of an almost textless page in which only images are presented (see page 26, line 14 to page 27, line 8 and Fig. 6 of the present specification). This effect cannot be achieved unless the specific automatic information filtering method, as explicitly recited in claim 1, is used.

In contrast, Hughes et al. completely fails to disclose any teaching for the use of the upper level URL derived from the top page URLs indicating inappropriate information, or the use of the automatic filtering based on words extracted from the information (page) indicated by the URL, as explicitly recited in claim 1. Hughes et al. only describes an undesirable content filtering using the primary filtering list and the URL keyword search, i.e., a detection of a prescribed word in the URL itself.

In fact, the quoted portions of Hughes et al., i.e. Abstract col. 2 lines 30-51; col. 2, line 30 to col. 5, line 21; col. 8, lines 44-50; col. 4, lines 50-67; col. 8, lines 17-50; col.10, line 40 to col. 11, line 27; cols. 4-5, lines 50-21; and col. 10, line 40-col. 11, line 64, do not even mention any derivation of an upper level URL from the top page URL or a list of such upper level URL, as well as any extraction of words from the information (page) indicated by the URL or the automatic filtering using such extracted words.

With respect to the argument by applicants that Hughes et al. failed to disclose any teaching for the use of upper level URL derived from the top page URLs indicating inappropriate information or the use of the automatic filtering based on words extracted from the information (page) indicated by the URL, the Examiner contended on page 7 of the Final Rejection that

“Hughes et al. teaches that “the final method of managing inappropriate material is ‘Content Filtering’. This involves scanning the actual material (not the URL) inbound to a net work from the Internet. Word lists and phrase pattern matching techniques are used to determine if the material is inappropriate or not. This process requires a great deal of computer processor time and power, slowing down Internet access and also making this a very expensive alternative.” (Hughes et al., col. 2, lines 17-26)”. But the Examiner omitted the next sentence from the quoted portion of Hughes et al., which reads “Furthermore, it is easily defeated by pictures, Java, or some other method of presenting words/content without the actual use of fonts.” How can Hughes et al. be considered an anticipatory teaching when the portion relied on by the Examiner is directly contrary to the teaching of applicants that the present invention provides method and apparatus for automatic information filtering which are capable of improving both precision and recall and accurately judging inappropriateness of the content even for a page that contains very few or no text information and only displays images, by utilizing an upper level URL of a URL given in a hierarchical structure.

The Examiner tacitly admits that Hughes et al. does not anticipate claim 1 by seeking to rely on Russell-Falla et al to support his contention on page 8 of the Final Rejection that “it would have been obvious to one of ordinary skill in the networking art to conclude that “the use of automatic filtering based on words extracted from the information (page) indicated by the URL” is well known feature in the networking art.” The attempt to incorporate an obviousness rejection into an anticipation rejection is not proper. The Russell-Falla patent was not part of the final rejection of claims 1-15 by the Examiner, which was solely anticipation under 35 U.S.C.

102 by Hughes et al. Furthermore, the shortcomings of Hughes et al. are not overcome by Russell-Falla et al.

Similarly, the Examiner seeks to rely on the disclosure of Humes to support his contention on page 8 of the Final Rejection that “it would have been obvious to one of ordinary skill in the networking art to conclude that “the use of the upper level URL derived from the top page URLs indicating inappropriate information is also well-known feature in the networking art.” The shortcomings of Hughes et al. are not overcome by Humes. Furthermore, the basis of the rejection of the appealed claims is anticipation under 35 U.S.C. 102. From decisions of the Federal Circuit, it is clear that strict, not substantial identity, is required for anticipation. Anticipation is established only if (1) all the elements of an invention, as stated in a patent claim, (2) are identically set forth, (3) in a single prior art reference. See, for example, Transclean Corp. v. Bridgewood Services, Inc., 62 USPQ2d 1865 (Fed. Cir. 2002); Sandt Technology, Ltd. v. Resco Metal and Plastics Corp., 60 USPQ2d 1091 (Fed. Cir. 2001); and Mehl/Biophile International Corp., 52 USPQ2d 1303 (Fed. Cir. 1999). It is incumbent upon the Examiner to identify wherein each and every facet of the claimed invention is disclosed in the applied reference. This has not been done here. Hughes et al. clearly does not anticipate claim 1, and the Examiner's rejection based on Hughes et al. is flawed. The rejection of claim 1 by the Examiner should be reversed.

The same argument also applies to the dependent claim 2, as well as corresponding apparatus claim 3 and 4 and corresponding medium claims 5 and 6.

Turning to claim 7, it should be noted that, in the claimed automatic information filtering method, the automatic learning using learning data containing both inappropriate information and appropriate information is carried out to obtain the word weights of words, and these word weights are stored and managed in correspondence to respective words. Then, the words contained in the entered information (page) are extracted and a total sum of the word weights of these extracted words is calculated, and whether the presentation of the entered information should be blocked or not is judged according to the calculated total sum of the words weights.

In contrast, Hughes et al. fails to disclose any teaching for the use of the word weights obtained by the automatic learning, or the use of the total sum of the words weights of the words extracted from the entered information (page) in judging the presentation of the entered information. As already mentioned above, Hughes et al. only describes an undesirable content filtering using the URL keyword search, i.e., a detection of a prescribed word in the URL itself.

In fact, the quoted portions of Hughes et al, i.e., col. 10, line 40 to col. 11, line 27 and col. 4, lines 50-67, do not even mention any learning of word weights or any use of such word weights or any use of such words weights in judging the presentation of the entered information. Col. 4 at lines 50-67 of Hughes et al. refers to URL Scanning and indicates that "URL Scanning looks for keywords, phrases, and pattern recognition within the URL as it is handled by the proxy server." Col. 10, line 40, et seq., indicates that "each and every proxy server log record is inspected and compared against the contents of the URL with a list of words, phrases and patterns. If the URL matches one or more of the key words or phrases or matches one of a list of specified patterns a notification is sent to the administrator." Nowhere in the portions of Hughes

et al. quoted by the Examiner is there a suggestion of “word weights” or the use of word weights in judging the presentation of the entered information.

Clearly, Hughes et al. does not anticipate claim 7, and the Examiner's rejection based on Hughes et al. is in error. The rejection of claim 7 by the Examiner should be reversed.

The same argument also applies to the claims 8 and 9, which depend from claim 7.

Since claim 10 was rejected by the Examiner employing the same logic as applied to claim 7, applicants submit that the arguments made as to claim 7 are also applicable as to claim 10. The rejection of claim 10 is in error and should be reversed.

Claims 11 and 12 depend from claim 10 and should be allowed together with claim 10 for the reasons stated above.

Claims 13-15, which were rejected as being corresponding computer readable medium claims of claims 7-9 under the same rationale as claims 7-9 should be allowed for the reasons expressed above with respect to claims 7.

9. Conclusion

Claims 1-15 are not anticipated by Hughes et al. Reversal of the rejection of claims 1-15 is earnestly solicited.

Respectfully submitted,

Dated: November 24, 2004

By Seymour Rothstein
Seymour Rothstein (Reg. No. 19,369)

CERTIFICATE OF MAILING

I hereby certify that this paper and its attachments are being deposited with the United States Postal Service on the date shown below with sufficient postage as First Class Mail in an envelope addressed to: Mail Stop APPEAL BRIEF-PATENTS, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450 on November 24, 2004.

Seymour Rothstein
Seymour Rothstein

APPENDIX 1

CLAIMS ON APPEAL

Claim 1 (original): A method of automatic information filtering for identifying inappropriate information among various information provided through Internet and blocking presentation of identified inappropriate information, comprising the steps of:

entering an HTML (HyperText Markup Language) information provided through the Internet;

judging whether a URL (Uniform Resource Locator) of said HTML information entered from the Internet is a top page URL or not, the top page URL being a URL ending with a prescribed character string defining according to a URL hierarchical structure by which each URL is constructed;

extracting words appearing in information indicated by the top page URL and carrying out an automatic filtering to judge whether said information indicated by the top page URL is inappropriate or not according to the words extracted from said information indicated by the top page URL, when said URL of said HTML information is the top page URL;

registering an upper level URL derived from the top page URL into an inappropriate upper level URL list and blocking presentation of said information indicated by the top page URL, when said information indicated by the top page URL is judged as inappropriate by the automatic filtering, the upper level URL being derived from the top page URL by keeping a character string constituting the top page URL only up to a rightmost slash;

comparing said URL of said HTML information with each URL registered in the inappropriate upper level URL list and judging whether there is any matching URL in the inappropriate upper level URL list when said URL of said HTML information is not the top page URL, and blocking presentation of information indicated by said URL of said HTML information when there is a matching URL in the inappropriate upper level URL list, the matching URL being one upper level URL whose character string is contained in said URL of said HTML information;

extracting words appearing in said information indicated by said URL of said HTML information, and carrying out the automatic filtering to judge whether said information indicated by said URL of said HTML information is inappropriate or not according to the words extracted from said information indicated by said URL of said HTML information, when there is no matching URL in the inappropriate upper level URL list; and

blocking presentation of said information indicated by said URL of said HTML information when said information indicated by said URL of said HTML information is judged as inappropriate by the automatic filtering.

Claim 2 (original): The method of claim 1, further comprising the steps of:
registering in advance URLs that provide inappropriate information in an inappropriate URL list; and

carrying out a third part rating based filtering for comparing said URL of said HTML information with each URL registered in the inappropriate URL list and judging whether there is any matching URL in the inappropriate URL list, and blocking presentation of said information indicated by said URL of said HTML information when there is a matching URL in the inappropriate URL list.

Claim 3 (original): An automatic information filtering apparatus for identifying inappropriate information among various information provided through Internet and blocking presentation of identified inappropriate information, comprising:

an input unit for entering an HTML (HyperText Markup Language) information provided through the Internet;

a top page URL judging unit for judging whether a URL (Uniform Resource Locator) of said HTML information entered from the Internet is a top page URL or not, the top page URL being a URL ending with a prescribed character string defined according to a URL hierarchical structure by which each URL is constructed;

a first automatic filtering unit for extracting words appearing in information indicated by the top page URL and carrying out an automatic filtering to judge whether said information

indicated by the top page URL is inappropriate or not according to the words extracted from said information indicated by the top page URL, when said URL of said HTML information is the top page URL;

an inappropriate upper level URL list registration unit for registering an upper level URL derived from the top page URL into an inappropriate upper level URL list and blocking presentation of said information indicated by the top page URL is judged as inappropriate by the automatic filtering, the upper level URL being derived from the top page URL by keeping a character string constituting the top page URL only up to a rightmost slash;

an inappropriate URL judging unit for comparing said URL of said HTML information with each URL registered in the inappropriate upper level URL list and judging whether there is any matching URL in the inappropriate upper level URL list when said URL of said HTML information is not the top page URL, and blocking presentation of information indicated by said URL of said HTML information when there is a matching URL in the inappropriate upper level URL list, the matching URL being one upper level URL whose character string is contained in said URL of said HTML information;

a second automatic filtering unit for extracting words appearing in said information indicated by said URL of said HTML information, and carrying out the automatic filtering to judge whether said information indicated by said URL of said HTML information is inappropriate or not according to the words extracted from said information indicated by said URL of said HTML information, when there is no matching URL in the inappropriate upper level URL list; and

an information presentation blocking unit for blocking presentation of said information indicated by said URL of said HTML information is judged as inappropriate by the automatic filtering.

Claim 4 (original): The apparatus of claim 3, further comprising:

an inappropriate URL list registration unit for registering in advance URLs that provide inappropriate information in an inappropriate URL list; and

a third party rating based filtering unit for carrying out a third part rating based filtering for comparing said URL of said HTML information with each URL registered in the inappropriate URL list and judging whether there is any matching URL in the inappropriate URL list, and blocking presentation of said information indicated by said URL of said HTML information when there is a matching URL in the inappropriate URL list.

Claim 5 (original): A computer usable medium having computer readable program codes embodied therein for causing a computer to function as an automatic information filtering apparatus for identifying inappropriate information among various information provided through Internet and blocking presentation of identified inappropriate information, the computer readable program codes include:

a first computer readable program code for causing said computer to enter an HTML (HyperText Markup Language) information provided through the Internet;

a second computer readable program code for causing said computer to judge whether a URL (Uniform Resource Locator) of said HTML information entered from the Internet is a top page URL or not, the top page URL being a URL ending with a prescribed character string defined according to the URL hierarchical structure by which each URL is constructed;

a third computer readable program code for causing said computer to extract words appearing in information indicated by the top page URL and carry out an automatic filtering to judge whether said information indicated by the top page URL is inappropriate or not according to the words extracted from said information indicated by the top page URL, when said URL of said HTML information is the top page URL;

a fourth computer readable program code for causing said computer to register an upper level URL derived from the top page URL into an inappropriate upper level URL list and block presentation of said information indicated by the top page URL, when said information indicated by the top page URL is judged as inappropriate by the automatic filtering, the upper level URL being derived from the top page URL by keeping a character string constituting the top page URL only up to a rightmost slash;

a fifth computer readable program code for causing said computer to compare said URL of said HTML information with each URL registered in the inappropriate upper level URL list and judge whether there is any matching URL in the inappropriate upper level URL list when said URL of said HTML information is not the top page URL, and block presentation of information indicated by said URL of said HTML information when there is a matching URL in the inappropriate upper level URL list, the matching URL being one upper level URL whose character string is contained in said URL of said HTML information;

a sixth computer readable program code for causing said computer to extract words appearing in said information indicated by said URL of said HTML information, and carry out the automatic filtering to judge whether said information indicated by said URL of said HTML information is inappropriate or not according to the words extracted from said information indicated by said URL of said HTML information, when there is no matching URL in the inappropriate upper level URL list; and

a seventh computer readable program code for causing said computer to block presentation of said information indicated by said URL of said HTML information when said information indicated by said URL of said HTML information is judged as inappropriate by the automatic filtering.

Claim 6 (original): The computer usable medium of claim 5, wherein the computer readable program codes further include:

an eighth computer readable program code for causing said computer to register in advance URLs that provide inappropriate information in an inappropriate URL list; and

a ninth computer readable program code for causing said computer to carry out a third part rating based filtering for comparing said URL of said HTML information with each URL registered in the inappropriate URL list and judging whether there is any matching URL in the inappropriate URL list, and blocking presentation of said information indicated by said URL of said HTML information when there is a matching URL in the appropriate URL list.

Claim 7 (original): A method of automatic filtering for identifying inappropriate information filtering for identifying inappropriate information among various information provided through Internet and blocking presentation of identified inappropriate information, comprising the steps of:

obtaining word weights of words to be used in judging whether presentation of each information should be blocked or not according to words contained in each information, by an automatic learning using learning data containing inappropriate information whose presentation should be blocked and appropriate information whose presentation should not be blocked;

storing and managing the word weights in correspondence to respective words in a form of a weighted word list; extracting words contained in information entered from the Internet; and

reading out the word weight for each word extracted from said information, from the weighted word list, calculating a total sum of the word weights of the words extracted from said information, and judging whether or not presentation of said information should be blocked or not according to the total sum.

Claim 8 (original): The method of claim 7, wherein the automatic learning is based on a linear discrimination function that can discriminate the inappropriate information and the appropriate information on a vector space.

Claim 9 (original): The method of claim 7, further comprising the steps of:
registering in advance URLs that provide inappropriate information in an inappropriate URL list; and

carrying out a third party rating based filtering for comparing said URL of said HTML information with each URL registered in the inappropriate URL list and judging whether there is any matching URL in the inappropriate URL list, the blocking presentation of said information indicated by said URL of said HTML information when there is a matching URL in the inappropriate URL list.

Claim 10 (original): An automatic information filtering apparatus for identifying inappropriate information among various information provided through Internet and blocking presentation of identified inappropriate information, comprising:

a word weight learning unit for obtaining word weights of words to be used in judging whether presentation of each information should be blocked or not according to words contained in each information, by an automatic learning using learning data containing inappropriate information whose presentation should be blocked and appropriate information whose presentation should not be blocked;

a weighted word list storing unit for storing and managing the word weights in correspondence to respective words in form of a weighted word list;

a word extraction unit for extracting words contained in information entered from the Internet; and

a judging unit for reading out the word weight for each word extracted from said information, from the weighted word list, calculating a total sum of the word weights of the words extracted from said information, and judging whether or not presentation of said information should be blocked or not according to the total sum.

Claim 11 (original): The apparatus of claim 19, wherein the automatic learning is based on a linear discrimination function that can discriminate the inappropriate information and the appropriate information on a vector space.

Claim 12 (original): The apparatus of claim 10, further comprising:
an inappropriate URL list registration unit for registering in advance URLs that provide inappropriate information in an inappropriate URL list; and
a third party rating based filtering unit for carrying out a third party rating based filtering for comparing said URL of said HTML information with each URL registered in the inappropriate URL list and judging whether there is any matching URL in the inappropriate URL list, and blocking presentation of said information indicated by said URL of said HTML information when there is a matching URL in the inappropriate URL list.

Claim 13 (original): A computer usable medium having computer readable program codes embodied therein for causing a computer to function as an automatic information filtering apparatus for identifying inappropriate information among various information provided through Internet and blocking presentation of identified inappropriate information, the computer readable program codes include:

a first computer readable program code for causing said computer to obtain word weights of words to be used in judging whether presentation of each information should be judging whether presentation of each information should be blocked or not according to words contained in each information, by an automatic learning using learning data containing inappropriate information whose presentation should be blocked and appropriate information whose presentation should not be blocked;

a second computer readable program code for causing said computer to store and manage the word weights in correspondence to respective words in a form of a weighted word list;

a third computer readable program code for causing said computer to extract words contained in information entered from the Internet; and

a fourth computer readable program code for causing said computer to read out the word weight for each word extracted from said information, from the weighted word list, calculate a total sum of the word weights of the words extracted from said information, and judge whether or not presentation of said information should be blocked or not according to the total sum.

Claim 14 (original): The computer usable medium of claim 13, wherein the automatic learning is based on a linear discrimination function that can discriminate the inappropriate information and the appropriate information on a vector space.

Claim 15 (original): The computer usable medium of claim 13, wherein the computer readable program codes further include:

a fifth computer readable program code for causing said computer to register in advance URLs that provide inappropriate information in an inappropriate URL list; and

a sixth computer readable program code for causing said computer to carry out a third part rating based filtering for comparing said URL of said HTML information with each URL registered in the inappropriate URL list and judging whether there is any matching URL in the inappropriate URL list, the blocking presentation of said information indicated by said URL of said HTML information when there is a matching URL in the inappropriate URL list.